DOCUMENT RESUME

ED 074 449 CS 000 398

AUTHOR Kennedy, John J.; And Others
TITLE The Effects of Induced "E" [Experimenter] Bias on the Reading of Instructions during a Behavioral Experiment.
SPONS AGENCY Ohio State Univ., Columbus. Coll. of Education.
PUB DATE Mar 73
NOTE 23p.; Paper presented at Annual Meeting of American Educational Research Association (New Orleans, Louisiana February 25-March 1, 1973)

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Bias; Expectation; *Experimental Psychology; *Oral Communication; Oral Reading; Reading; *Suprasegmentals; *Test Bias; *Testing; Testing Problems

ABSTRACT

Eight graduate students were used as experimenters in this study to assess the effects of experimentally induced experimenter outcome bias with respect to selected suprasegmental phenomena (pitch, stress, and terminal intonation) emitted by the experimenters during the instruction reading phase of a behavioral experiment. Experimenters were led to expect differential conditioning performance from the subjects prior to the experiment. The subjects, 24 undergraduate students, were informed as to the nature of the experiment and were instructed to provide results which were compatible with the experimenters' expectations. The experimenters' directions to the subjects were audio-taped and rated by trained judges. Analyses of data produced several marginal findings but generally failed to demonstrate that bias is communicated through systematic variations in the measured language features. It was noted, however, that the response measures analyzed in this study were more specific than measures examined in earlier research and that the failure to uncover anticipated relationships suggests the study direct attention to an anomaly or inconsistency in the chain of logic underlying a substantive phenomenon which can serve to prompt further exploration and clarification. (Directions to students and a consensus sheet are appended) (HS)

ED 074449

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
John J. Kennedy
Victor M. Rentel
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER

THE EFFECTS OF INDUCED E BIAS ON THE READING OF INSTRUCTIONS DURING A BEHAVIORAL EXPERIMENT*

John J. Kennedy+
Victor M. Rentel
Robert Griffin

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
Robert Griffin
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER

Experimenter bias (E-bias), the apparent ability of Es to unintentionally influence the behavior of subjects (Ss) in a direction consonant with their a priori expectations, constitutes one of the most active areas for research and writing in both social psychology (Rosenthal, 1966; Friedman, 1967; Jung, 1971) and education (Rosenthal & Jacobson, 1968). Although the viability of the E-bias phenomenon has been challenged in several contexual environments (Barber & Silver, 1968; Kennedy, 1969; Elashoff & Snow, 1971), the most recent comprehensive review of the literature in question (Rosenthal, 1969) clearly documented the existence of E expectancy effects in several laboratory settings. Moreover, Rosenthal's review pointed to the direction-reading phase of the typical behavioral experiment as the place where the communication of bias most likely occurs and further suggested that auditory cues (as opposed to bodily or facial cues) constituted the most promising general area for research into the still uncharted realm of E-bias mediation.

---

*Paper presented at the meeting of the American Educational Research Association, New Orleans, March 1973. This research was supported by a small grant from The College of Education's Research Committee, The Ohio State University.

+John J. Kennedy is affiliated with the faculty of Development, The Ohio State University, Columbus, Ohio. Victor M. Rentel is affiliated with the Reading Faculty in Early and Middle Childhood, The Ohio State University, Columbus, Ohio. Robert Griffin is with the Department of Modern and Classical Languages, Western Michigan University, Kalamazoo, Michigan.

CS 000 398

Some notion of the importance of the auditory cues emitted during instruction reading, relative to the mediation of E bias, can be inferred from the Rosenthal, Friedman and Kurland (1966) study in which male Es, but not female Es, who read instructions more rapidly and with less accuracy tended to obtain more hypothesis confirming data. Moreover, Adair and Epstein (1967) have shown that E expectancy effects can occur when Ss are exposed to only the audio-taped instructions of biased Es. But the most convincing evidence, thus far, has been provided by Duncan and his associates. Specifically, Duncan and Rosenthal (1968), in an exploratory study, found a significant relationship ($r=.72$, $p<.01$) between a developed measure of vocal emphasis and Ss' behavior on a photo-rating task. The measure of vocal emphasis was derived from a complex mix of suprasegmental phomemes (stress, pitch, terminal juncture) and paralinguistic phenomena (intensity, pitch level, voice openness, pitch range, tempo, etc.) taken ex post facto from the recorded instructions of Es who in earlier research had achieved outcomes consistent with their a priori expectations. In this descriptive research, key word-pairs were identified in Es' directions to Ss and the vocal emphases associated with word-pairs were rated by the principal investigator. As noted, a relationship between vocal emphasis and Ss' performance was examined, but comparisons between Es who had been induced experimentally to hold different outcome expectations were not made and as such was cited by the authors as a major limitation.

In a subsequent study, Duncan, Rosenberg, and Finkelstein (1969) exposed Ss to audio-taped directions with predetermined differential verbal emphasis scores, the latter being determined as in the earlier Duncan and Rosenthal (1968) study. Again, vocal emphasis was found to relate significantly to Ss' performance suggesting that Es expectations can be communicated to Ss through subtle manipulations of voice. But again, comparisons between Es who had been indoctrinated to expect different outcomes were not made.

The purpose of this study was to explore the effects of experimentally induced E expectancy with respect to the number and nature of selected supra-segmental features emitted by Es during the reading of experimental instructions. Specifically, this research sought to identify variations in suprasegmental features such as stress, terminal juncture (pause), and terminal intonation as a function of Es' outcome expectation. In addition to the principal independant variable (i.e., outcome expectancy set), the type of pre-experimental training provided Es, and the sex of Es were also considered. Type of training was of obvious interest because intensive and thorough training of Es prior to their participation in research has long been advanced as a method of minimizing bias and experimental error. Further, lack of standardization in instruction reading has been shown to be related to the communication of E-bias (Rosenthal, Friedman & Kurland, 1966). On several occasions, E's sex has also been shown to be a correlate of experimental outcome, although in the majority of instances, observed sex effects have been difficult to meaningfully interpret and predict (Rosenthal, 1966, Chapter 4).

METHOD

Experimenters

Eight graduate students, four males and four females, served as Es. Initially, the Es, selected in part because they were relatively unfamiliar with the literature in E expectancy and verbal conditioning, were only informed that they were to attempt a replication of the classic Taffel experiment (Taffel, 1955) in verbal conditioning. Es were paid ($20) for their services.

E Bias indoctrination. Subsequent to selection, Es were subjected to an indoctrination, designed to lead Es to expect positive results (significant verbal conditioning) from some of their Ss (n=5) and negative results (nonconditioning) from others (n=5). Briefly, the outcome expectancy set was established for each E during a briefing session with the principal investigator. During this session, Es were informed that a similar replication of the Taffel experiment had been conducted previously by the principal investigator, but results were extremely variable and inconclusive. However, the experience gained from that investigation led the present research team to believe that they could now explain the variable results. The key factor, the Es were told, appeared

to be the manner in which Ss were recruited. Ss, who were "pressured into" participating, will, on the average, fail to condition while volunteer Ss, who generally will want to please, will demonstrate significant conditioning effects. The Es were also told that to "balance the experimental design," each E would be assigned five volunteer Ss (conditioners) and five required Ss (nonconditioners). Thus, through the briefing, the eight Es were led to believe that the purpose of the experiment was to confirm a hypothesis calling for differential conditionability as a function of Ss' recruitment. Parenthethically, similar procedures have been previously used with success relative to establishing pre-experimental outcome expectancy (see Kennedy, Cook, & Brewer, 1970).

Training in direction reading. Es were exposed to one of two distinct training regimens prior to participation. Specifically, two male and two female Es were subjected to a rigorous training program which emphasized, in general, the importance of procedural standardization and, in particular the importance of reading experimental directions accurately and uniformly. With respect to the reading of directions, the four Es participated in several pilot sessions after which, in the Es presence, audio-taped recordings of direction reading were reviewed and severely critiqued by the principal investigator. On the other hand, the remaining four Es (two males and two females) were subjected to a more relaxed and casual training program. (Probably typical of the training of Es in many behavioral studies.) Casual training simply consisted of several non-evaluative role playing sessions and one pilot session with a "live" S.

Subject-Accomplices

The Ss were 24 undergraduate students, largely sophomores, who in this experiment were not Ss in the traditional sense but rather accomplices. Specifically, Ss were informed as to the nature of the experiment. Furthermore, they were instructed to provide results which were compatible to Es' expectations, especially during the initial sessions. For example, if an E were running his or her first "volunteer" S (i.e., his or her first anticipated conditioner), the S-accomplice would provide the E with data that clearly indicated that conditioning had occured. Conversely, if E were running his or her first "nonvolunteer" S, the S-accomplice would provide negative results. During subsequent successions, S-accomplices were instructed to give Es results that, on the average, tended to conform to the E's expectation.

The Experimental Task and Procedures

The Taffel task. To both enhance the face validity of the research effort and to provide a creditable opportunity to obtain specimens of Es direction reading behavior, each E engaged ten Ss (accomplices) in a standard Taffel-type verbal conditioning task. Essentially, the materials for the task consist of eighty five-by-eight stimulus cards each containing a different past tense verb typed (primer) in the center of the card. Below each verb, also in primer type, were six pronouns: I, WE, YOU, HE, SHE, and THEY. The order of pronoun presentation was randomly determined for each card.

The task was performed in a small office containing a table, two chairs, and an audio-tape recorder. Prior to an experimental session, the E was given

the name of the S and whether the S volunteered (and thus could be expected to condition) or was pressured into participating (and thus could be expected to provide negative results). Es then recorded this S-status information on a prepared tabulating sheet which Es believed constituted an important record for the experiment. At the moment of an S's appearance, the E started the tape recorder. Ss were then introduced to the E by one of the investigators and were requested to sit at the table opposite the E. The stack of stimulus cards, which was shuffled before each session , was placed before the S and the E, then, read verbatum the procedural instructions.

Experimental instructions. Essentially, the instructions requested Ss to make up a sentence for each card but the sentence must begin with one of the presented pronouns and it must incorporate, somewhere, the verb contained on the card. The instructions, which consisted of 18 sentences and are presented in Appendix A, were designed not only to direct Ss but also to provide Es with many opportunities to exercise variations in oral stress (the relative force with which a syllable is uttered), pause (hesitation or delay), and terminal intonation (the manner of terminating an utterance, prior to a pause, by altering the pitch). An examination of these directions will reveal that they contain many pronouns, which conceivably could provide Es with opportunities to unconsciously communicate to Ss that self-reference pronouns (Ior WE) are the desired operants. Further, there is even an opportunity (see sentence 13) for Es to personally select one of the relevent pronouns to use as an example to Ss.

For purposes of analysis, thirty words in the directions were identified by these investigators as words which could be differentially treated by Es to communicate outcome expectancy. Generally, these words were pronouns (e.g., "I", "You", "Us" etc.) or words that conceivably could be used to "tip-off" a S as to the desired self-reference operant (e.g., "personal", "Yours"). These words, termed key words are circled in Appendix A.

Experimental sessions. During actual sessions, Es were verbally mute for the first 20 trials (cards). Es attempted to reinforce the emittance of self-reference pronouns (i.e., I or WE) during trials 21 through 80 by responding with a neutrally toned "mm-hmm" immediately after each sentence emitted by a S that began with the operant. Es recorded Ss' responses manually on prepared tabulation sheets. Tabulation sheets were given to one of the investigators immediately upon the termination of a session and the degree to which S responded to the E's attempted reinforcement was jointly determined. During initial sessions, particularly, the investigator made a point of demonstrating to the E that the S's (accomplice) performance was in accordance with his or her a priori hypothesis. This procedure was designed to maintain and reinforce differential E expectations established earlier during the briefing session with the principal investigator (see Rosenthal, 1966, Chapter 12).

As a final check on the effectiveness of efforts to establish differential outcome expectations as a function of supposed S recruitment, the eight participating Es were interviewed subsequent to the experiment. The results of these audio-taped interviews suggested that seven of the eight Es still believed that the purpose of the study was to replicate Taffle's classic experiment and that S recruitment status was a significant determinant as to whether or not significant conditioning would occur. In one instance, the E admitted that she felt as if there might be a "secret purpose" to the experiment because she was so carefully "observed" prior to and after experimental sessions, however, she was unable to specify what the secret purpose was.

Experimental Variables and Design

Four independent variables pertaining to Es defined the generic design of this study. These variables were:

1. Type of training (T)--a fixed active variable consisting of two levels:
   T_1 rigorous training (previously descirbed)
   T_2 casual training (previously described)

2. Sex of E (X)--a fixed, assigned variable consisting of two levels:
   X_1 male Es
   X_2 female Es

3. Individual Es (Es)--an assumed random variable consisting of two Es nested within each of four Training-Sex (TX) combinations

4. E outcome expectancy (B)--a within E variable consisting of two levels:
   B_1 positive (anticipated successful conditioning)
   B_2 negative (anticipated nonconditioning)

Ten S-accomplices were assigned to each E. Five of the Ss were labeled as "volunteers", and hence conditioners, and five were labeled "required Ss" and thus were not expected to produce positive results. The order or sequence of assignment was done randomly. Independent variables were organized such that T and X were crossed, levels of Es were nested within TX combinations, and levels of B comprised a within E variable. A diagram of experimental group arrangements presented by Figure 1.

Pre-Experimental Training	Sex of Es	Individual Es	Positive	Negative
Rigorous	Male	E_1	S_{11111} $S_{21111}\cdots S_{51111}$	S_{11112} $S_{2112}\cdots S_{51112}$
		E_2	S_{11121} $S_{21121}\ldots S_{51121}$	$S_{11122}\ldots$
	Female	E_3	S_{11211} $S_{21211}\cdots$	
		E_4	$S_{11221}\ldots$	
Casual	Male	E_5	$S_{12111}\cdots$	
		E_6	S_{12121} $S_{22121}\cdots$	
	Female	E_7	⋮	
		E_8	S_{12221} $S_{22221}\cdots S_{52221}$	S_{12222} $S_{22222}\cdots$

Figure 1 Schema of Experimental group arrangements

Dependent variables that were essentially interval in nature were subjected to a univariate analysis of variance. Relative to the ANOVA, the outcome expectancy variable (B) was treated as a within Ss or repeated measurements variable. In addition, the five Ss comprising an expectancy group (positive or negative) were treated as a within-group or repeated measurements variable. The rational for this was that dependent measures were not derived from the performance of individual Ss (accomplices) but rather from features of spoken directions emitted by an E and the same E read directions to sets of five Ss whom, it is assumed, the E believed would condition or fail to condition depending on the group in question. Therefore, for interval appearing data, a 2 X 2 X 2 X 2 X 5 mixed model ANOVA was employed. For dependent variables which were only amenable to nominal classification, simplier two-way chi-square analyses were performed.

The Rating of Spoken Direction

Raters were recruited from a graduate level course in applied English phonology which had just treated in depth segmental and suprasegmental features of the English Language. Initially, seven members of this class expressed interest in the assignment and were invited to participate in an orientation session. After an explanation of procedures for identifying and scoring suprasegmental features, students were provided with several speech samples to evaluate. Although there was near perfect agreement among the seven students during this initial session, the three students whose analysis of samples appeared to be most accurate and consistent were selected as raters. These three raters then participated in four more hours of training supervised by the third investigator prior to the actual scoring sessions.

All three raters were present together to score each audio-tape. Audio-tapes of E's direction reading were presented at random to the raters who were blind to E's expectancy. The tape recording of each reading was stopped at the end of every sentence and the raters scored intonation contours, stress, and pause features within each sentence unit. At the end of the entire reading, raters compared their results and then replayed any section of the taped specimen for which there was doubt or disagreement. After problems were resolved, the final evaluations were noted on a consensus sheet (see Appendix B). This procedure was observed for each of the eighty taped readings.

Dependent Variables

Because of the exploratory nature of this study, an attempt was made to identify as many potentially relevant measures or indicies of prosodic features as possible. Presented below are the measures which were generated for study and subjected to analysis in this study. The majority of these measures were subjected to a univatiate analysis, as opposed to MANOVA, because the intent was not to document group differences relative to a profile of dependent measures in linear combinations but rather to identify a specific prosodic feature or features which might serve as a mediator of E bias.*

*In linguistic science the study of phonetic intonation and the systematic role of pitch contours, as well as levels, remains as yet not fully elaborated at least within a framework of syntactic or phonological theory. (Chomsky and Halle, 1968, p.ix). Nor is the work in paralanguage sufficiently advanced that

(footnote continued on next page)

	Code	Principle Analysis
1. Measures of Stress or Accent		
a) Total number of stress markings (both primary and secondary) on each consensus sheet	TS	ANOVA
b) Number of stress markings for each of the 18 sentences	S-S	ANOVA
c) Type of stress associated with each of the 30 key words in the directions (i.e., no stress, secondary stress, primary stress)	S-K	X^2
d) Ratio of the number of key words stressed to the total number of stress markings per consensus sheet	K/TS	ANOVA
2. Measures of Pause		
a) Total number of pause markings on each consensus sheet	TP	ANOVA
b) Number of pause markings for each of the 18 sentences (i.e., number of utterances per sentence unit)	P-S	ANOVA
c) Type of pause associated with each of the 30 key words in the directions (i.e., no pause adjacent to key word; pause just prior to key word; pause just subsequent to key word; pause before and after key word	P-K	X^2
d) Ratio of the number of pauses associated with key words to the total number of pause markings per consensus sheet	K/TP	ANOVA
3. Measures of Terminal Intonation		
a) Total number of terminal intonations by type for each consensus sheet: (1) level (-)	L-I	ANOVA

(footnote continued from page 7) variables derived from clear theoretical antecendents may be explored. As these fundamental linguistic structures become known, research attempting to describe the relationship between expectancy and effect will be greatly facilitated. The variables of stress, pitch, and juncture as identified in this study must be considered crude in this respect. As will be seen, we have attempted to examine the rate and ratios of occurrence of certain obligatory features of stress, pitch, and juncture but have no underlying theoretical justification for our descriptions, much less our predictions.

(2)	falling (^)	F-I	ANOVA
(3)	rising (v)	R-I	ANOVA

b) Ratio of intonation type to total number of expressed intonations per consensus sheet

(1)	level to total	L/TI	ANOVA
(2)	Falling to total	F/TI	ANOVA
(3)	rising to total	R/TI	ANOVA

c) Type of intonation pattern for each sentence unit. Examples:

(1) — (i.e., one utterence--level terminal intonation)	X^2
(2) ^	X^2
(3) v (i.e., one utterance--rising intonation)	X^2
(4)-- (i.e., two utterances--both level)	X^2
(5)-^	X^2
(6)-^	X^2
⋮	
(14)--^ (i.e., three utterances, level-level-falling)	X^2
⋮	
(30)^v^	X^2
⋮	
(39)vvv	X^2

4. Other Measures

a)	Es personal choice of pronoun in the example sentence (Sentence 18). Eg., "I", "HE" etc.	PN	X^2
b)	Type of stress associated with selected pronoun	S-PN	X^2

RESULTS

Analysis of Measures of Stress

The analysis of measures of stress involved, in part, performing a 2 X 2 X 2 X 2 X 5 univate ANOVA on: a) total stress (TS), b) the number of stresses observed for each sentence ($S\text{-}S_1$, $S\text{-}S_2$, ...$S\text{-}S_{18}$), and c) the ratio of key words stressed to total stress (k/TS). Presenting a summary of all analysis would be prohibitive, therefore a representative sampling of findings are given below. Specifically, Table 1 shows the mean performance observed for measures of TS, $S\text{-}S_3$ (Sentence 3), $S\text{-}S_8$ (Sentence 8), and K/TS by E outcome expectation (bias), type of training, and sex of E. Relative to the most interesting effect, the difference between overall means of outcome expectancy groups, an examination of Table 1 suggested that only minor group differences were pre-

sent.

Insert Table 1 here

The analysis of variance on the measures presented in Table 1 is summarized by Table 2. Aside from depicting the dimensions of the design, Table 2 also revealed that few, if any, meaningful results materialized. For example, the analysis of variance of Total Stress (TS) failed to produce one significant effect at the .05 level, the level of significance established for this research. While Table 2 shows that the main effects for Es' Sex were significant ($F=9.57$; $df = 1/4$; $P<.05$) for number of stresses associated with the Third Sentence in the directions, the fact that female Es emitted more stresses in connection with this particular sentence than did males appears to have little theoretical or methodological importance. Similarly, the significant effects for Es /TX observed in both the analysis of stresses for Sentence 8 and the K/TS ratios simply indicated that significant variability occured among individual Es with respect to these measures. The significant Es B/TX source also seen in the analysis of stress markings for Sentence 8 simply suggests a first order interaction between individual Es and outcome expectancy but these interactions must be interpreted within the context of Training-Sex combinations thus also rendering them uninteresting. And, as implied above, the summarized analyses were typical of the remaining analyses not shown.

Insert Table 2 here

The type of stress given to each of the 30 designated key words in the directions were analyzed by chi-square tests for homogeniety. For each key word, a 2 X 2 table was constructed where key words read to expected conditioners and expected nonconditioners were further catagorized on the basis of whether the key word was stressed (either primary or secondary) or not stressed during delivery. Table 3 presents two typical analyses, the analysis of the word PERSONAL in the fifth sentence of the directions and the analysis of YOU appearing in sentence 15 of the directions. As can be seen, both analyses failed to produce a significant result. The remaining 28 analyses, not shown, also failed to yield significance. In sum, these negative findings suggest that Es, who have undergone an indoctrination relative to outcome expectancy, do not consciously or unconsciously communicate elements of this expectation to Ss by differentially stressing what appeared to be "key words" (or cue words) in the directions.

Insert Table 3 here

Analysis of Measures of Pause

Essentially the same procedures and methods of analysis were employed for the measures of pause cited in the METHODS section. The 2 X 2 X 2 X 2 X 5

ANOVA on total number of pause markings (TP) produced a significant F for Es nested within TX combinations ($F = 6.08$; $df = 4/16$; $p < .005$) and a significant F ($F = 3.26$; $df = 4/16$; $p < .05$) for the first-order interaction between Es and Expectancy within TX combinations (Es B/TX). However, the former possessed no substantive relevance while subsequent examination of the latter significant effect revealed only that one E (rigorous-male) paused totally more often for expected conditioners than nonconditioners while another E (casual-female) paused more often for nonconditioners than suspected conditioners. Overall pause behavior, therefore, does not appear to be a promising vehicle relative to the communication of E bias.

The 18 univariate analyses performed on the number of pauses associated with each sentence ($P\text{-}S_1$, $P\text{-}S_2$,...$P\text{-}S_{18}$) resulted in a total of ten significant effects at .05. However, as before, few of these effects were relevant to the purposes of this study. One possible exception was an observed first-order interaction between Sex and Expectancy (XB) for the Sixth Sentence in the directions ($F = 4.45$; $df = 1/4$; $p < .12$). Subsequent post-hoc tests revealed a tendency for female Es to pause more frequently in reading Sentence 6 to expected conditioners (average of 1.45 pauses) than to expected nonconditioners (1.25). Again, however, it must be remembered that this finding did not achieve the significance criterion established for this research.

The analysis of variance of K/TP ratios produced only a significant Es /TX effect ($p < .05$) suggesting significant variability among individual Es independent of Training or Sex conditions.

To determine if there was a difference in the number of pauses associated with key words as a function of the Es assumed outcome expectation, two-variable chi-square analyses were performed. Specifically, a X^2 was performed for each of the 30 words where the first classification variable consisted of the two categories of expectancy (positive-negative) and the second variable consisted of the categories: a) no pause adjacent to key word, b) pauses prior to key word, c) pauses subsequent to key word, and d) pauses both prior to and after key word. In all instances, significance was not observed.

Analysis of Measures of Terminal Intonation

As previously mentioned, the intonation given to the last spoken syllable prior to each pause was rated as either a) rising (such as occurs as the end of some interrogatives, b) falling (such as that associated with some simple declaratives, or c) level (such as occurs in conjunction with some declaratives and expletives). For each of the three types of terminal intonation (i.e., L-1, F-1, and R-1), an ANOVA was performed on the total number of intonation markings per consensus sheet. Each of the three analyses produced a significant F for the Es /TX source, again indicating that the most consistent result over the many analyses was that individual Es manifested significant differences independent of the conditions manipulated in the study or E's sex. The remaining handful of significant effects were equally uninteresting.

Ratios of type of terminal intonation to total number of recorded intonation marking per consensus sheet (i.e., L/TI, F/TI, and R/TI) were also subjected to the analysis of variance and also failed to produce findings that were of interest.

Initially, it was felt that the analysis of intonation patterns associated with each of the sentences in the directions would prove to be most productive. The analysis was performed by recording the frequency with which each pattern occured for each sentence and then testing (X^2) for differences in pattern pro-

file between expectancy conditions. For example, with respect to the Fourth Sentence in the directions, in the majority of instances this sentence was verbally read as one utterance (continuous discourse initiated and terminated with a pause) with falling intonation (^) associated with the last word (i.e., the verb printed on the first card which served as an example--see Appendix B). However a scattering of additional intonation patterns were also observed, e.g., two utterances of the form - ^ (level-falling), and the form - v (level-rising) etc.. The chi-square analysis for Sentence 4 is given by Table 4.

Insert Table 4 here

The resultant chi-square statistic ($X^2 = 3.40$) fell short of significance at the .05 level. Similar analyses were conducted for the remaining 17 sentences with essentially the same general outcome: nonsignificance. Thus, what at first had appeared to be a promising approach also failed to produce results concerning the specific mediation of E bias.

Finally, chi-square analyses on the pronoun selected by Es during the example sentence (Sentence 18) and the type of stress given to the selected pronoun by outcome expectation were conducted. There was no evidence to suggest either the differential selection of pronouns or the differential stress afforded pronouns between expectancy conditions.

SUMMARY AND DISCUSSION

The purpose of this study was to assess the effects of experimentally induced E outcome bias with respect to selected suprasegmental phenomena emitted by Es during the instruction reading phase of a behavioral experiment. Es were led to expect differential conditioning performance from Ss prior to the experiment. Es' directions to Ss were audio-taped recorded and ratings on pitch, stress, and terminal intonation were provided by trained judges. Analyses of data produced several marginal findings; but generally, this research failed to demonstrate that bias is communicated through systematic variations in the measured language features.

It is generally recognized that a study which fails to advance significant results has scholarly merit only to the extent to which it accomplishes at least one of the following: a) the study describes novel methodological approaches which have not produced the desired or anticipated result and, therefore, contributes to knowledge concerning the methodology associated with a substantive

area; b) the negative results emanating from the study sharply contradict previous results and thus recreate a "problem situation" to be addressed by future research; or c) the study directs attention to an anomoly or inconsistency in the chain of logic underlying a substantive phenomenon which can serve to prompt further exploration and clarification.

With respect to the first accomplishment, new methodology, it should be noted that the response measures analyzed in the present study (i.e., stress, pause, and terminal intonation) were more specific than measures examined in the previously cited research. For example, the Duncan et. al. studies (1968, 1969) employed a measure, termed vocal emphasis, which was highly catholic. Specifically, the vocal emphasis measure, which reflected the differential emphasis given to key or cue words in a limited segment of instructions, was generated by one rater, albeit an expert rater, who made his determinations on the basis of a composite of several suprasegmental phenomena and a variety of paralanguage features. Acknowledging the fact that significance was achieved in both of the Duncan studies, the inability to observe significance in the present study suggests the possibility that the manipulation of specific suprasegmental features has little psychological meaning relative to the communication of outcome expectancy to Ss. However, when such features are combined and viewed as a more complex and global entity (e.g., vocal emphasis), meaning, and hence significance, is achieved. In short, the "whole" may not be distinguishable by an analysis of its constituent parts.

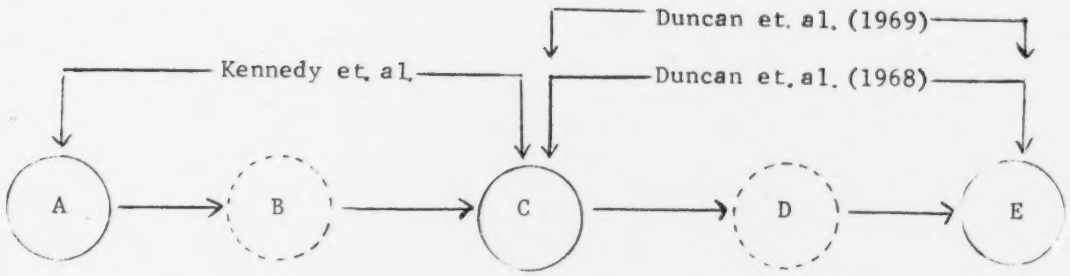
The second cited function, that of challenging existing or emerging findings, clearly is not served by the present effort. Aside from the fact that the specific language variables examined in the present study differed from that used by Duncan and associates, major differences still remain in the basic design of these studies. In general, the variable of greatest interest in the Duncan

studies was the measure of vocal emphasis associated with individual Es. The dependent variable in these studies was Ss response on Rosenthal's (1966) photo-rating task. The present study, on the other hand, attempted to manipulate the outcome expectations of Es(the independent variable)while subsequent measures of suprasegmental phenomena taken on individual Es constituted the dependent variables. In sum, any attempt to challenge Duncan's findings is unwarrented because the Duncan studies and the present study do not examine the same substantive domains.

The fact that the described study examined conditions antecedent to the differential use of suprasegmental features by Es (to influence Ss), and failed to uncover anticipated relationships between these conditions (the procedural act of biasing Es), suggests that the third function of nonsignificant studies might be served by this research. Consider the diagram below (Figure 2) which depicts, in temporal and causal sequence, several major events leading to S-compliance with Es' expectation. It has been noted that a significant relationship between events (C) and (E) ($r = .73$, $p < .01$) has been described (Duncan & Rosenthal, 1968) and experimentally varified (Duncan et. al., 1969)* The present study, however, searched for significant effects between (A) and (C), assuming that Es adopted the expectancy set implemented in (A) but none were found. Thus, this study suggests that gaps in our understanding of the mediation of E bias still exist, at least in terms of the simplest chain of causal events depicted in Figure 2, and that further systematic exploration is still needed.

---

*With respect to the first Duncan study (1968), a correlation between Es' expectancy indoctrination (A) and Ss performance (E) was computed ($r = .60$, $p < .04$), however, only Es were included who were known to have "biased" their Ss responses in the predictable direction which clearly vitiates the significance of this result.



(A.) The procedural act of indoctrinating Es to expect certain outcomes, or differential outcomes, as a function of S characteristics. (A practice common to many research efforts into E-bias)

(B.) Es' assimilation of the expectancy set. (A mentalistic event which at present cannot be directly measured)

(C.) Es' unintended manipulation of suprasegmental and paralanguage features during direction reading to communicate outcome expectation to Ss.

(D.) Ss' preconscious or conscious interpretation of Es' expectation. (A mentalistic event)

(E.) Ss' behavioral compliance in the direction of preceived expectation.

Figure 2 Hypothesized events prior to S's compliance to E's outcome expectation.

More precisely, we assume that any sort of communication, bias or otherwise, is a structures entity. These structured embody and specify the relationship between form and conceptual import. Our failure to specify a set or subset of these structures having clear theoretical antecedents constituted a major weakness in this study. Unfortunately modern linguistic theory sheds little light on either the surface structure of suprasegmentals or their underlying phonological, syntactic, and semantic representations. Lieberman (1965) has argued that the

perception of intonation can be accounted for by a generative model but not by systems such as that given by Trager-Smith (1957). Stockwell (1964) illustrated two morphophonemic expansion rules that would place pitch, terminal juncture and primary phrase stress, which, taken with the Halle-Stevens (1959) "analysis by synthesis" model suggest an interesting possibility for future research.

REFERENCES

Adair, J. G. & Epstein, J. Verbal cues in the mediation of experimenter bias. Psychological Reports, 1968, 22, 1045-1053.

Barber, T. X., & Silver, M. J. Fact, fiction, and the experimenter bias effect. Psychological Bulletin, 1968, 70, (Part 2, Monogr. Suppl.).

Chomsky, N. and Halle, M. The sound pattern of English. New York: Harper & Row, 1968.

Duncan, S., Rosenberg, M. J., & Finkelstein, J. The paralanguage of experimenter bias. Sociometry, 1969, 32, 207-219.

Duncan, S. & Rosenthal, R. Vocal emphasis in experimenters' instruction reading as unintended determinant of subjects' responses. Language and Speech, 1968, 11, 20-26.

Eloshoff, J. D. & Snow, R. E. Pygmalion Reconsidered. Worthington, Ohio: Charles A. Jones, 1971.

Friedman, N. The social nature of psychological research. New York: Basic Books, 1967.

Halle, M. and Stevens, K. N. Analysis by synthesis. In W. Wathon & L.E. Woods (Eds.), Proceedings of the seminar on speech comprehension and processing. Bedford, Mass: Air Force Cambridge Research Laboratories, 1959.

Jung, J. The experimenter's dilemma. New York: Harper & Row, 1971.

Kennedy, J. J. Experimenter outcome bias in verbal conditioning: A failure to detect the Rosenthal effect. Psychological Reports, 1969, 25, 495-500.

Kennedy, J. J., Cook, P. A., & Brewer, R. R. The effects of three experimenter variables in verbal conditioning research. Journal of Social Psychology, 1970, 81, 167-175.

Lieberman, P. On the acoustic basis of the perception of intonation by linguists. Word, 1965, 21, 40-54.

Rosenthal, R. Experimenter effects in behavioral research. New York: Appleton-Century-Crofts, 1966.

Rosenthal, R. Interpersonal expectations. In R. Rosenthal & R. L. Rosnow (Eds.), Artifact in Behavioral Research. New York: Academic Press, 1969, pp. 181-277.

Rosenthal, R., Rriedman, N., & Kurland, D. Instruction-reading behavior of the experimenter as an unintended determinant of experimental results. Journal of Experimental Research in Personality, 1966, 1, 221-226.

Rosenthal, R. & Jacobson, L. Pygmalion in the classroom. New York: Holt, Rinehart, & Winston, 1968.

Stockwell, R. P., The place of intonation in a generative grammar of English. Language, 1964, 36, 360-367.

Taffel, C. Anxiety and the conditioning of verbal behavior. Journal of Abnormal and Social Psychology, 1955, 51, 496-501.

Trager, G.L. & Smith, H. L. An outline of English structure, Washington, D.C., 1957.

TABLE 1

MEANS OF A REPRESENTATIVE SAMPLE OF PAUSE MEASURES

Outcome Expectancy	Training	Es' Sex	TS	$S\text{-}S_3$	$S\text{-}S_8$	K/Ts
Positive (expected Ss would condition	Rigorous	M	49.60	1.40	3.90	.46
		F	62.40	1.90	5.40	.60
	Casual	M	55.00	1.20	4.30	.58
		F	55.20	1.70	4.00	.67
Negative (did not expect Ss to condition)	Rigorous	M	53.50	1.50	3.60	.71
		F	60.40	2.50	4.70	.52
	Casual	M	55.90	1.00	4.50	.46
		F	60.40	2.00	4.70	.66
Overall Positive			55.55	1.55	4.40	.58
Overall Negative			57.55	1.75	4.38	.59

TABLE 2

ANOVA OF REPRESENTATIVE STRESS MEASURES

Source	df	TS		S - S_3		S - S_8		K/TS	
		MS	F	MS	F	MS	F	MS	F
Between E_s:	7								
Training (T)	1	.4500	.01	2.4500	2.09	.0125	.00	.0092	.04
Sex of E (X)	1	744.2000	5.41	11.2500	9.57*	7.8125	1.34	.0708	.37
TX	1	281.2500	2.04	.0000	.00	9.1125	1.57	.1496	.79
E_s/TX[a]	4	137.6750	2.50	1.1750	1.66	5.7875	4.01*	.1893	3.47*
Within E_s:	72								
Expectancy (B)	1	80.0000	.96	.8000	1.52	.0125	.00	.0036	.08
TB	1	22.0500	.26	.4500	.86	4.5125	1.12	.1264	2.82
XB	1	3.2000	.04	1.2500	2.38	.0125	.00	.0594	1.33
TXB	1	130.0500	1.57	.0000	.00	1.0125	.25	.2576	5.75
E_sB/TX	4	82.9750	2.07	.5250	.85	4.0375	4.58*	.0448	.88
S_s Within B (S)	4	51.8250	.94	.7063	1.00	.9500	.66	.0197	.36
TS	4	7.6375	.14	.2938	.42	.1375	.10	.0811	1.49
XS	4	78.2000	1.42	.9063	1.28	.2500	.17	.0331	.61
TXS	4	76.1875	1.39	.4688	.66	.7375	.51	.0732	1.34
E_s S/TX	16	54.9875	1.38	.7063	1.14	1.4438	1.63	.0545	1.07
BS	4	70.8125	1.77	.6437	1.04	1.9500	2.21	.0217	.42
TBS	4	38.3000	.95	.6063	.98	.5125	.58	.0458	.90
XBS	4	90.2625	2.26	.7188	1.16	1.2000	1.36	.0697	1.37
TXBS	4	103.6750	2.60	.1562	.25	1.2625	1.43	.0637	1.25
E_sBS/TX	16	39.9125		.6188		.8813		.0511	
Total	79								

*$p < .05$

TABLE 3

CHI-SQUARE ANALYSIS OF TWO OF

THE 30 KEY WORDS IN THE DIRECTIONS

"PERSONAL" (sentence no. 5)

Type of Pause

	None	Primary	Total
Expected Conditioners	29 (31.5)[a]	11 (8.5)	40
Expected Nonconditioners	34 (31.5)	6 (8.5)	40
Totals:	63	17	80

$X^2 = 1.87$ ($p < .20$)

"YOU" (sentence no. 15)

Type of Pause

	None	Primary or Secondary	Total
Expected Conditioners	27 (26.5)	13 (13.5)	40
Expected Nonconditioners	26 (26.5)	14 (13.5)	40
Totals:	53	27	80

$X^2 = .06$ ($p < .85$)

a - Theoretical frequencies are contained within parentheses.

TABLE 4

ANALYSIS OF INTONATION PATTERN BY OUTCOME EXPECTANCY FOR THE FOURTH SENTENCE IN THE DIRECTIONS

Intonation Pattern

	One utterance falling (∧)	Others (—∧, —∨ etc.)	Totals
Positive Expectancy	21 (25)	19 (15)	40
Negative Expectancy	29 (25)	11 (15)	40
	50	30	80

$X^2 = 3.41$

APPENDIX A

DIRECTIONS TO SUBJECTS

BEFORE YOU* IS A STACK OF 5 by 9 CARDS. YOU WILL SEE THAT EACH CARD CONTAINS A STANDARD VERB -- AND BELOW EACH VERB YOU WILL SEE A SERIES OF PERSONAL PRONOUNS. WHY DON'T WE LOOK AT THE FIRST CARD. YOU CAN SEE THAT THE VERB FOR THIS CARD IS (indicate the verb). AND NOTE THE SERIES OF PRONOUNS (read the six pronouns). AS YOU CAN SEE PERSONAL PRONOUNS INDICATE A RELATIONSHIP NOT ONLY TO SELF BUT TO OTHERS AS WELL.

WHEN I TELL YOU TO BEGIN, YOU ARE TO MAKE UP A SENTENCE FOR EACH CARD. MORE SPECIFICALLY, YOU ARE TO MAKE UP SENTENCES WHICH USE THE VERB FROM THE CENTER OF THE CARD. BUT I WANT YOU TO SELECT ONE OF THE PERSONAL PRONOUNS TO START EACH SENTENCE. IN OTHER WORDS, SELECT ANY ONE OF THE SIX PERSONAL PRONOUNS TO BEGIN THE SENTENCE ... THE CHOICE IS YOURS. YOU ARE FREE TO CREATE A SENTENCE SPONTANEOUSLY UTILIZING THE VERB IN ANY POSITION OR IN ANY FORM THAT YOU CHOOSE. YOU NEED NOT CONNECT YOUR SENTENCES THROUGH SOME MANNER OF NARRATIVE SEQUENCE.

FOR EXAMPLE, FOR THE CARD THAT I HAVE JUST SHOWN YOU, YOU MIGHT SAY (give an example). — EXAMPLE PRONOUN

YOU ARE TO SELECT CARDS -- IN ORDER -- FROM THE PILE -- MAKE UP A SENTENCE USING THE MATERIAL ON THAT CARD -- AND THEN, PLACE THE CARDS FACE DOWN HERE (point). YOU SHOULD CONTINUE TO MAKE UP SENTENCES UNTIL YOU HAVE RUN THROUGH ALL THE CARDS.

ARE YOU READY?...FINE. LET US BEGIN WITH THE FIRST CARD.

* Circled words were designated key words by these investigators.

APPENDIX B

Tape _____
Scorer _____
Date _____

CONSENSUS SHEET

Directions to Subjects

Volunteer-Conditioning Study

SENTENCE NO.		Pause No.1			Pause No.2			Pause No.3		
		ACCENT		INT.	ACCENT		INT.	ACCENT		INT
		P	S		P	S		P	S	
1.	BEFORE YOU IS A STACK OF 5 BY 9 CARDS.									
2.	YOU WILL SEE THAT EACH CARD CONTAINS A STANDARD VERB-- AND BELOW EACH VERB YOU WILL SEE A SERIES OF PERSONAL PRONOUNS.									
3.	WHY DON'T WE LOOK AT THE FIRST CARD.									
4.	YOU CAN SEE THAT THE VERB FOR THIS CARD IS (indicate the verb).									
5.	AND NOTE THE SERIES OF PRONOUNS (I, you, he, she, it, we, they).									
6.	AS YOU CAN SEE PERSONAL PRONOUNS INDICATE A RELATIONSHIP NOT ONLY TO SELF BUT TO OTHERS AS WELL.									
7.	WHEN I TELL YOU TO BEGIN, YOU ARE TO MAKE UP A SENTENCE FOR EACH CARD.									
8.	MORE SPECIFICALLY, YOU ARE TO MAKE UP SENTENCES WHICH USE THE VERB FROM THE CENTER OF THE CARD.									
9.	BUT I WANT YOU TO SELECT ONE OF THE PERSONAL PRONOUNS TO START EACH SENTENCE.									
10.	IN OTHER WORDS, SELECT ANY ONE OF THE SIX PERSONAL PRONOUNS TO BEGIN THE SENTENCE. . . THE CHOICE IS YOURS.									
11.	YOU ARE FREE TO CREATE A SENTENCE SPONTANEOUSLY UTILIZING THE VERB IN ANY POSITION OR IN ANY FORM THAT YOU CHOOSE.									
12.	YOU NEED NOT CONNECT YOUR SENTENCES THROUGH SOME MANNER OF NARRATIVE SEQUENCE.									
13.	FOR EXAMPLE, FOR THE CARD THAT I HAVE JUST SHOWN YOU, YOU MIGHT SAY _____.									

Volunteer-Conditioning Study
November 1970

	Pause No.1			Pause No.2			Pause No.3		
	ACCENT		INT.	ACCENT		INT.	ACCENT		INT.
	P	S		P	S		P	S	
YOU ARE TO SELECT CARDS -- IN ORDER -- FROM THE PILE --									
14. MAKE UP A SENTENCE USING THE MATERIAL ON THAT CARD --									
AND THEN PLACE THE CARDS FACE DOWN HERE.									
15. YOU SHOULD CONTINUE TO MAKE UP SENTENCES									
UNTIL YOU HAVE RUN THROUGH ALL THE CARDS.									
16. ARE YOU READY?. . . FINE. 17.									
18. LET US BEGIN WITH THE FIRST CARD.									